Sequential Estimation of Quantiles *

Jong-Suk R. Lee[†], *Donald McNickle*[‡] *and Krzysztof Pawlikowski*[†]

[†]*Department of Computer Science and* [‡]*Department of Management*

University of Canterbury

Christchurch, New Zealand

Abstract

Quantiles are convenient measures of the entire range of values of simulation outputs. However, unlike the mean and standard deviation, the observations have to be stored since calculation of quantiles requires several passes through the data. Thus, quantile estimation(QE) requires a large amount of computer storage and computation time. Several approaches for estimating quantiles in RS(regenerative simulation) and non-RS, which can avoid the difficulties of QE, have been proposed in [Igl76], [Sei82b], and [JC85].

In this report, we implemented these three approaches known as: *linear* QE, *batching* QE, and *spectral* P^2 QE are studied in the context of sequential steady-state simulation. Numerical results of coverage analysis of these three QE approaches are presented.

Keywords : sequential simulation, quantile estimation, regenerative, batch means, spectral analysis, P^2 algorithm

1 Introduction

In simulating a stochastic system, such as a queueing or inventory system, the simulator is frequently more concerned with the extreme performance of the system than with the long run average. As opposed to averages, quantiles can account for extreme behaviour of the system. Quantiles are convenient measures of the entire range of values of simulation outputs. Analysts find quantiles particularly useful in estimating reasonable capacities for facilities, comparing the overall performance of alternative designs or establishing minimum standards. Therefore, from a practical point of view, the problem of estimating quantiles is quite important.

Let $X_1, ..., X_n, ..., X_N$ be a sample of i.i.d random variables from a continuous distribution $F_X(x)$ with probability density function $f_X(x)$. For $0 < p < 1$, let

$$x_p = \inf \{x : F_X(x) \geq p\} = F_X^{-1}(p),$$

*Technical Report TR-COSC 05/98

where $F_X^{-1}(p)$ is the inverse of $F_X(x)$ with derivative $1/f_X(x_p)$. The quantity x_p is the pth quantile of $F_X(x)$.

Let $X_{(1)} \leq ... \leq X_{(n)} \leq ... \ X_{(N)}$ be the order statistics corresponding to the sample. The usual non-parametric point estimator of x_p is the pth sample quantile

$$\hat{x}_p = X_{(\lfloor Np+1 \rfloor)}, \quad 0 < p < 1$$

where $\lfloor z \rfloor$ denotes the integral part of z.

The problem with using $\hat{x}_p$ as an estimator is that the sample sizes required for adequate precision are prohibitively large. Both sorting times and memory sizes are then unrealistic. A measure of the inflation of sample size over the independence case has been investigated by Blomqvist. Unlike the mean and standard deviation, the observations have to be stored since calculation of quantiles requires several passes through the data. Thus, quantile estimation(QE) requires a large amount of computer storage, and very long runs for securing the credibility of the final results. It makes the amount of computation time very large. For example, extreme quantiles of the M/M/1/∞ queueing system with traffic intensity $\rho = 0.9$ requires a sample size of roughly 500,000 customers to estimate the 0.99 quantile of the waiting time distribution to within plus or minus 10% accuracy. For the 0.999 quantile, the required sample size is approximately 2,3000,000. Clearly, storing and sorting the entire sequence is impractical in such a cases. Actually, to produce an order-statistic point estimate of x_p requires storing the largest $(1-p)/N$ values of the sequence. However, this ordering must be dynamically maintained as the sequence is generated, a computationally expensive operation and an additional storage is also required to estimate the variance. This above results is derived from Table 8 of Blomqvist by Heidelberger and Lewis in [HL84].

Several approaches for estimating quantiles in RS(regenerative simulation) and non-RS, which can avoid the above difficulties, have been proposed in [Igl76], [Sei82b], and [JC85]. These three approaches are originally developed for traditional(non sequential) procedure.

Our motivation is finding the robust estimator of quantiles in sequential steady-state simulation. Thus, in this report we discuss these three approaches known as *linear* QE, *batching* QE for RS, and *spectral* P^2 QE for non-RS, in the context of sequential steady-state simulation. In Section 2 detailed sequential QE approaches of RS, which are *linear* QE and *batching* QE, are discussed.In Section 3 detailed sequential QE approach of non-RS, which is *spectral* P^2 QE, is also discussed. In Section 4 the numerical results of coverage analysis of three QE approaches are presented and conclusions are in Section 5.

2 Sequential QE Approaches of RS

The regenerative method(RM) of simulation, first suggested by Cox and Smith, for analysis of observations collected from a regenerative process $\{X(t) : t \geq 0\}$ has been systematically developed by a number of authors. The regenerative approach is motivated by the fact that many stochastic systems have the property of "starting afresh probabilistically" from time to time. The central idea of the RM is to exploit the fact that, when $\{X(t) : t \geq 0\}$ is a regenerative process, random variables between successive regeneration points are independent and identically distributed(i.i.d.) thus it can circumvent the autocorrelation problem in estimates.

Iglehart[Igl76], Moore[Moo80], Seila[Sei82b], and Heidelberger and Lewis[HL84] have given special methods for processes $\{X_n\}$ with regenerative structure, i.e. processes for which there exist random time points at which the process restarts probabilistically. An example is the waiting time process $\{W_n\}$ in the M/M/1/∞ queueing system which regenerates every time a customer arrives to find the queue empty, so that the waiting time of that customer is zero.

Detailed comparisons of Iglehart, Seila and Moore's approaches for QEs in fixed sample size approach are in [Sei82a]. Three methods mentioned differ significantly. Each has advantages and disadvantages, and the appropriate method will depend on the specific application. The summary of three methods' comparisons is as following Table 1:

Table 1: Comparisons of Three Methods

Methods	**Statistical Precision**	**Computational Efficiency**	**Memory Efficiency**
Iglehart	Moderate	High	Moderate
Seila	Moderate	High	Moderate
Moore	High	Low	Low

These three QE approaches for RS use fixed-sample size analysis method even though sequential analysis of simulation output is generally accepted as the most efficient way for securing representativeness of samples of collected observations. Also, there are no consideration of variance reduction and produces a single quantile. In this paper, we consider two approaches, which are *Iglehart*'s (we will call *linear*) and *Seila*'s (we will call *batching*) methods for sequential QE because *Moore*'s approach didn't consider memory and computing time efficiency as we can see the Table 1.

Among the few possible criterion for stopping the simulation, probably the most commonly used one is based on the relative half width of the confidence interval at a given confidence level (1-α) defined as the ratio

$$\epsilon = \frac{\Delta_x}{\bar{X}(n)} \qquad 0 < \epsilon < 1; \tag{1}$$

where $\bar{X}(n)$ is the estimation of the mean μ_x of an analyzed process from the sequence of collected observations $x_1, x_2, \cdots, x_n$ and Δ_x is the half width of the confidence interval for the estimator. It is well known that if observations $x_1, x_2, \cdots, x_n$ can be regarded as realizations of independent and normally distributed random variables $X_1, X_2, \cdots, X_n$, then

$$\Delta_x = t_{n-1,1-\alpha/2}\hat{\sigma}\left[\bar{X}(n)\right],$$

where

$$\hat{\sigma}\left[\bar{X}(n)\right] = \sum_{i=1}^{n} \frac{\{x_i - \bar{X}(n)\}^2}{n(n-1)}$$

is the (unbiased) estimator of the variance of $\bar{X}(n)$, and $t_{n-1,1-\alpha/2}$ is the (1-α) quantile of the t-distribution with (n-1) degrees of freedom.

The ratio of Equation(1) is called the *relative precision of the confidence interval*. The simulation experiment is stopped at the first checkpoint for which $\epsilon \le \epsilon_{max}$, where ϵ_{max} is the required limit relative precision of the results at the 100(1-α)% confidence level, $0 < \epsilon_{max} < 1$.

2.1 Sequential QE using *Linear* Approach

The *Linear* approach has originally developed for fixed sample size simulation by Iglehart [Igl76]. In this paper, we modified the *linear* approach for sequential QE. First of all, *linear* approach for sequential QE requires to first specify a grid of h+1 points $a_0 < a_1 < a_2 < ... < a_h$ so that all observations lie between a_0 and a_h. In this report, we used 21 grid points spaced 0.2 units which is reasonable for QE (but not for extreme QE) because all theoretical QE values of M/M/1/∞ queueing system are in grids. Next, this method estimates the cumulative distribution function only at grid points. Then, it interpolates linearly between these estimates to find the quantile estimate until the steady state parameter has been estimated with the required relative precision.

The sample quantile, $Q_n(p)(= x)$, based on n RC(regenerative cycle)s would be defined by the relation

$$Q_n(p) = inf\{x : F_n(x) \ge p\}; 0 < p < 1.$$

Next consider how the function F_n and the sample quantiles $Q_n(p)$ would be constructed in the course of a simulation experiment. First suppose the regenerative process has a discrete state space. Assume the state space is the integers $\{0, 1, \cdots, N\}$. Then in the course of simulating the n RCs we would cumulate the time spent in each of the states. If $w_n(i), i = 0, 1, \cdots, N$, is the time spent in state i in n RCs, then F_n would jump an amount equal to $w_n(i)/\beta_n$ at state i where $\beta_n : n \ge 1$ is the increasing sequence of regeneration time. The sample quantiles after n RCs would then be computing by taking

$$Q_n(p) = min\{k \ge 0 : w_n(0) + \cdots + w_n(k) \ge p\beta_n\}.$$

Variance, $\sigma^2(Q_n(p))$, would be to estimate $\sigma^2([Q_n(p)])$ and $\sigma^2([Q_n(p)]+1)$ using Equation(2) and then linearly interpolate. The estimate $\hat{s}^2(n)$ corresponds to what is called the classical estimator.

$$\sigma^2([x]) = \sigma^2\{Y_1([x])\} - 2F([x])cov(Y_1([x]), a_1) + F^2([x])\sigma^2(a_1). \quad (2)$$

In Equation(2), $\sigma^2\{Y_1([x])\}$, $cov(Y_1([x]), a_1)$, and $\sigma^2(a_1)$ are estimated by the standard sample variances and covariance. $F([x])$ is estimated by $\sum_{i=1}^{n} Y_i([x])/\beta_n$

where Y_i is the sum of observations collected during the ith RC. The derivative $F'[Q(p)]$ is estimated by $W_n([Q_n(p)]+1)/\beta_n \equiv \hat{f}(n)$, the slope of F_n at the point $Q_n(p)$ and $\bar{a}(n)$ is estimated by $(1/n)\sum_{i=1}^{n} a_i$ where a_i is the length of the ith RC or, eqivalently, the number of observations collected during the RC i.

$100(1-\alpha)\%$ confidence interval for the quantile, $Q_n(p)$, is given by

$$Q_n(p) \pm z_{1-\alpha/2}\hat{s}/\bar{a}\hat{f}n^{\frac{1}{2}}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution.

2.2 Sequential QE using *Batching* Approach

Batching approach also has been originally developed for fixed sample size simulation by Seila [Sei82b]. In this report, we modified the *batching* approach for sequential QE. First, the *batching* approach for sequential QE groups data from the RCs into batches and uses sample quantiles computed from the batches as a set of independent, identically distributed observations. Before applying this method, the analysts must specify the batch size (the number of RCs in a batch) and in this report we considered 50 batch size.

Batching method groups m cycles in each batch. Within each batch, a quantile estimate is computed using the jackknifed sample quantile until the steady state parameter has been estimated with the required relative precision.

Batching method incorporates a two-fold jackknife in order to reduce bias as follows: Assume that m is even, and let $Q_{m/2,2i-1}$ and $Q_{m/2,2i}$ be the sample quantiles computed from the first $m/2$ and second $m/2$ cycles in the ith batch. The jackknifed batch quantile is

$$J(\hat{Q}_{m,i}) = 2\hat{Q}_{m,i} - \frac{1}{2}(\hat{Q}_{m/2,2i-1} + \hat{Q}_{m/2,2i}).$$

The sequence $\{J(\hat{Q}_{m,i}), i = 1, 2, ..., r\}$ consists of i.i.d. random variables. Letting $J(\bar{Q}_{m,r})$ and $\hat{\sigma}^2(J(\bar{Q}_{m,r}))$ denote the sample mean and variance:

$$J(\bar{Q}_{m,r}) = \frac{1}{r}\sum_{i=1}^{r} J(\hat{Q}_{m,i});$$

and

$$\hat{\sigma}^2(J(\bar{Q}_{m,r})) = \frac{1}{r-1}\sum_{i=1}^{r}(J(\hat{Q}_{m,i}) - J(\bar{Q}_{m,r}))^2.$$

$100(1-\alpha)\%$ confidence interval for the quantile, $J(\bar{Q}_{m,r})$, is given by

$$J(\bar{Q}_{m,r}) \pm t_{1-\alpha/2,r-1}\hat{\sigma}(J(\bar{Q}_{m,r}))/\sqrt{r}$$

where $t_{1-\alpha/2,r-1}$ is the $(1-\alpha/2)$ quantile of the student's t-distribution with $r-1$ degrees of freedom.

3 Sequential QE Approaches of Non-RS

A QE approach to overcome the typical difficulties of QE in non-RS has been proposed by Jain & Chlamtac [JC85]. This QE approach is based on P^2(*Piecewise-Parabolic*) formula (see *Appendix I*). QE using P^2 algorithm originally developed for fixed-sample-size procedure and non-RS.

The P^2 algorithm consists of maintaining five markers: the minimum, the $p/2$-, p-, and $(1+p)/2$-quantiles, and the maximum. The markers are numbered 1-5. Markers 2 and 4 are also called middle markers because they are midway between the p-quantile(marker 3) and the extremes. The y value (height) of each marker in Figure 7 is equal to the corresponding quantile value, and its x value is equal to the number of observations that are less than or equal to the marker. The marker heights are the current estimates of the quantiles, and these estimates are updated after every observation. As a new observation comes in, it is compared with the markers, and all markers higher than the observation are moved one position to the right. If a maker is off to the left or right of its ideal location by more than one, then the y and x values are adjusted using a P^2 formula. Pseudocodes of P^2 algorithm and an example calculation using P^2 algorithm are in [JC85].

QE using P^2 algorithm solves the storage problem by calculating quantiles using a piecewise-parabolic formula dynamically as the observations are generated. The observations are not stored, instead, a few statistical counters are maintained which help refine the estimate. Therefore, QE using P^2 algorithm has a very small storage requirement regardless of the number of observations and a small computing time because no sorting required.

Extended P^2 approach for QE, which is extended version of Jain and Chlamtac's approach, has been proposed by Raatikainen [Raa87] and this *Extended* P^2 approach simultaneously estimates several quantiles without storing and sorting the observations in fixed-sample-size procedure. Sequential procedure for simultaneous estimation of several quantiles in non-RS has also been proposed in [Raa90]. This sequential approach uses the *extended* P^2 algorithm to estimate the quantiles and the variances of the quantile estimates are estimated using SA/HW method (Spectral Analysis in its version proposed by Heidelberger and Welch [HW81]). This approach used the random length of uniform distribution in $[1025, 2048]$ for deciding the length of the initial transient period.

3.1 Sequential QE using *Spectral* P^2 Approach

In this report, we are considering sequential procedure for a single QE in non-RS and we will call *Spectral* P^2 Approach. P^2 algorithm proposed by Jain & Chlamtac was used for QE and the variances of the quantile estimates are estimated using SA/HW method. However, the method for detecting the initial transient period used here are based on those described in [Paw90]. The sequential procedure of [Paw90] for detecting the initial transient period is based on a stationarity test proposed by Schruben et al. [SST83]. It is used to test the hypothesis that a sufficient number of initial transient data has been discarded. As in any statistical test, the value of a chosen statistic calculated from the tested sequence of observations is compared with the corresponding value from a standard sequence, and the decision about rejection or acceptance of the hypothesis is taken at an assumed significance level α_t, $0 < \alpha_t < 1$. The significance level can be regarded as the probability of erroneously rejecting the hypothesis that the tested process is stationary.

Many heuristic rules are discussed in [Paw90], and to get a first approximation for the truncation point, a heuristic rule proposed by Fishman [Fis73], which is *The initial transient period is over after* n_0 *observations if the time series* $x_1, x_2, \cdots, x_{n_0}$ *crosses the mean* $\bar{X}(n_0)$ k *times*, is used for detecting the initial transient period. In Gafarian et al. [GAM78], $k = 25$ was recommended for M/M/1/∞ queueing systems so we assume $k = 25$ for simulation experiments. Detailed pseudocodes of the sequential procedure for detecting the initial transient period using Fishman's heuristic rule are in [Paw90]. Detailed methods and pseudocodes of the sequential procedure for testing the required precision of results using SA/HW method are also in [Paw90].

The p-quantile which is calculated by the P^2 algorithm (detailed procedure of this method in [JC85]) would be defined by Q_p. The P^2 formula for the heights, which is quantile value, is actually an approximation of the inverse of the empirical cumulative distribution function, $\hat{F}^{-1}(y) = ay^2 + by + c$. As observations N become large, the variance of quantile, Q_p, can be approximated by

$$\sigma^2(Q_p) = p(0; Q_p)/N\hat{f}(Q_p)^2,$$

where $p(0; Q_p)$ is the stationary spectral density which is estimated using the SA/HW method applied to the sequence $\{I_i(Q_p), i = 1, \cdots, N\}$ and $\hat{f}(Q_p)$ be the stationary density function which can be approximated by $\hat{f}(Q_p) = (b + 2a\hat{F}(Q_p))^{-1}$ in the neighbourhood of Q_p.

$100(1-\alpha)\%$ confidence interval for the quantile, Q_p, is given by

$$Q_p \pm t_{1-\alpha/2, r-1}\hat{\sigma}(Q_p)/\sqrt{r}$$

where $t_{1-\alpha/2,r-1}$ is the $(1 - \alpha/2)$ quantile of the student's t-distribution with $r - 1$ degrees of freedom.

4 Numerical Results

Implementations of the *batching* QE approach and the *linear* QE approach in RS and *Spectral* P^2 Approach in non-RS for analysing sequential QE of output data have been discussed in the previous section.

The robustness of any method can be usually measured by the coverage of confidence intervals, defined as the proportion $\hat{p}$ with which the number of the final confidence interval $(\hat{p} - \Delta, \hat{p} + \Delta)$ contains the true value p. An estimator of variance $\hat{\sigma}^2$ used for determining the confidence interval of the point estimate is considered as valid, i.e. producing valid $100(1-\alpha)\%$ confidence intervals of the point estimate, if the upper bound of the confidence interval of the point estimate $\hat{p}$ equals at least $(1-\alpha)$ [SL79]. Coverage analysis, however, is to analytically tractable systems, since the theoretical value of the interesting parameter has to be known. Because of this reason, in this report we estimated quantiles of response times of M/M/1/∞ queueing system.

All numerical results in this report were obtained by stopping simulation experiments when the final steady-state results reached a required precision of 5% or less, at the 0.95 confidence level and 200 or more bad confidence intervals (to secure representativeness in the analysed data) had been collected. All results were also filtered of strangely short simulation runs to secure the statistical properties of interval estimators after 200 bad confidence intervals are collected. The filtering of short simulation runs has improved estimates of coverage for sequential QE in RS except a case. The results of traditional fixed sample size approach with 200 replications are shown in Figure 2, Figure 4, and Figure 6.

Sequential coverage analysis (using F approximation ([LMP98], [LMP99]), which gives much narrower confidence intervals,) for sequential QE approaches in RS and non-RS on a single processor under MRIP (Multiple Replications In Parallel) scenario of AKAROA ([PYM94] and [EPM99]) are experimented at 0.9 quantile.

Figure 1 and Figure 3 depict the results obtained from *Linear* QE and *Batching* QE approaches in RS, respectively and Figure 5 is from *Spectral*P^2 QE approach in Non-RS. As we can see, *batching* QE approach in RS gives much stable coverage except traffic intensity 0.9.

5 Conclusions

For large number of observations, QE becomes impractical to store and sort the entire sequences. Physical memory limitations of computers used for QE make large numbers of replications impossible, and in others, the shuffling of virtual memory pages slow down the simulation considerably. *Linear* and *batching* QE approaches for RS and P^2 QE approach for non-RS can resolve the problems related with QE but these are not originally developed for sequential simulation.

Run length control of simulation is very important as the most efficient way for securing simulation results statistically. Sequential stopping rules, which control the relative width of an estimated confidence interval, can be used in conjunction with RS and non-RS. Therefore, we proposed three QE approaches for RS and non-RS in sequential steady-state simulation: *linear* QE, *batching* QE, and *spectral* P^2 QE.

If QEs of RS and non-RS are performed using multiple processors in MRIP scenario for example using simulation package AKAROA, it could reduce the computation time additionally.

Many other aspects for sequential QE in RS and non-RS will have to be carefully studied and tested with a number of different simulation models before these procedures can be safly used in simulation practice.

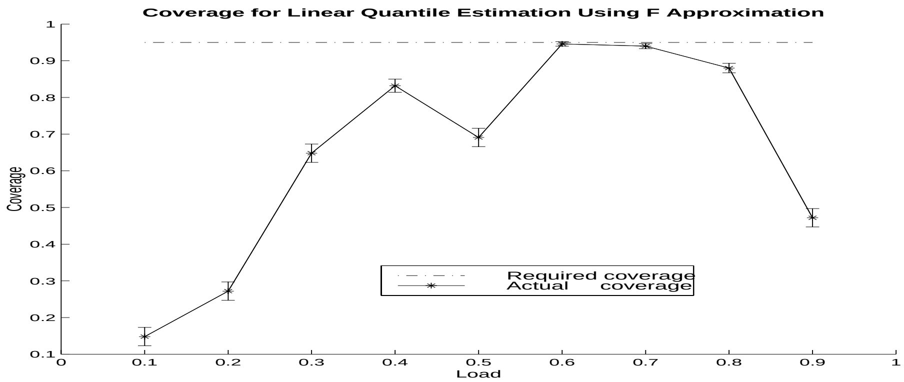

Figure 1: Coverage Analysis of *Linear* QE in RS Using F Approximation in M/M/1/∞ Queueing System ($P = 1$ & Sequential Analysis)

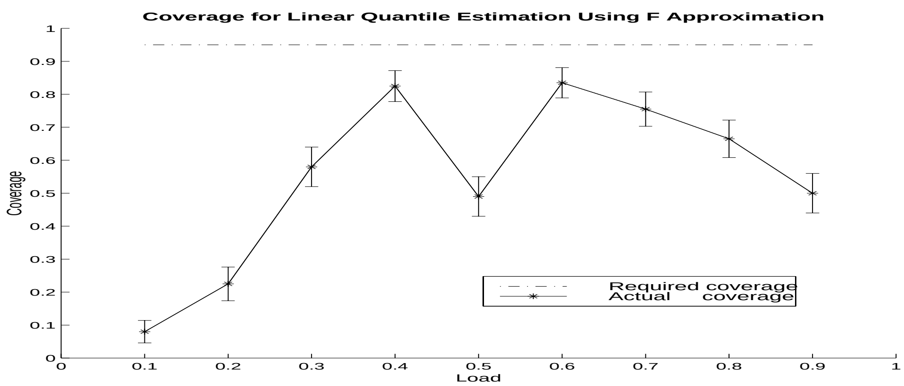

Figure 2: Coverage Analysis of *Linear* QE in RS Using F Approximation in M/M/1/∞ Queueing System ($P = 1$ & Fixed Sample Size of 200 Replications)

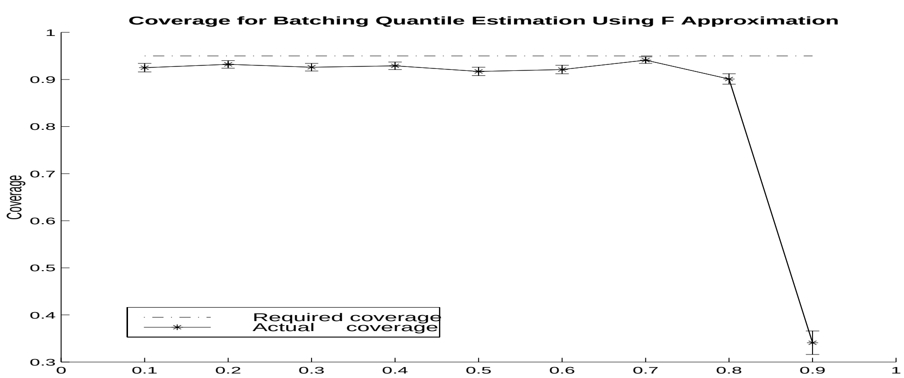

Figure 3: Coverage Analysis of *Batching* QE in RS Using F Approximation in M/M/1/∞ Queueing System ($P = 1$ & Sequential Analysis)

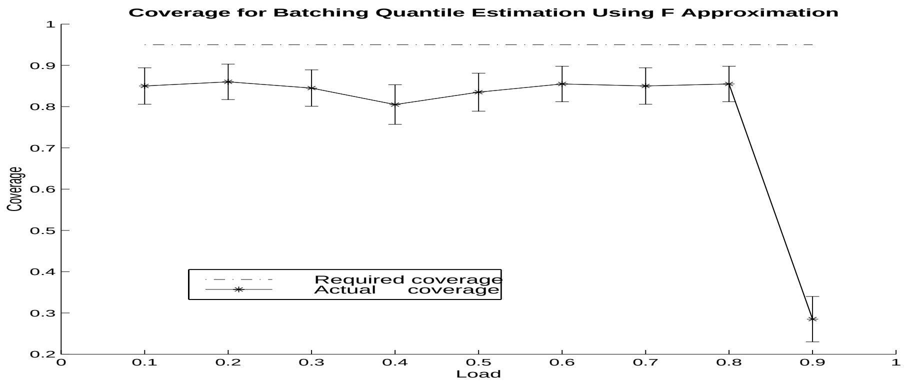

Figure 4: Coverage Analysis of *Batching* QE in RS Using F Approximation in M/M/1/∞ Queueing System ($P = 1$ & Fixed Sample Size of 200 Replications)

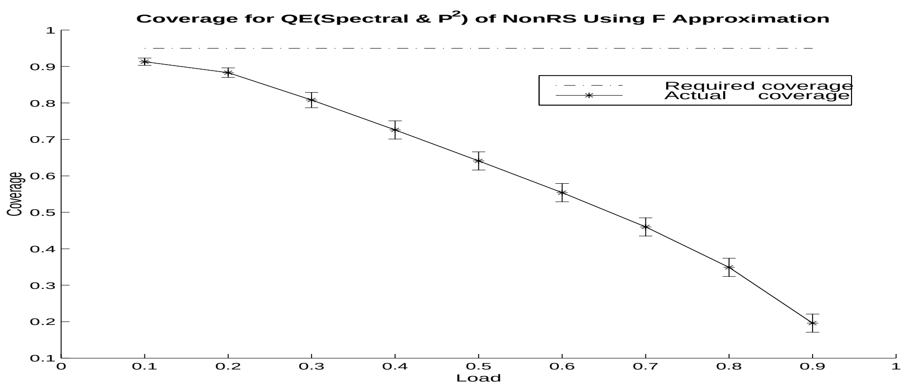

Figure 5: Coverage Analysis of *Spectral* P^2 QE in Non-RS Using F Approximation in M/M/1/∞ Queueing System ($P = 1$ & Sequential Analysis)

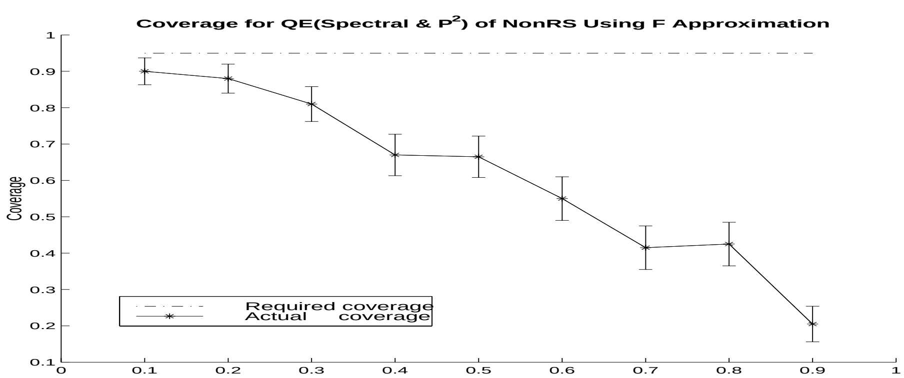

Figure 6: Coverage Analysis of *Spectral* P^2 QE in Non-RS Using F Approximation in M/M/1/∞ Queueing System ($P = 1$ & Fixed Sample Size of 200 Replications)

Appendix I: *Derivation of the P^2 Formula*

As shown in Figure 7, the P^2 formula assumes that the curve passing through (n_{i-1}, q_{i-1}), (n_i, q_i), and (n_{i+1}, q_{i+1}) is a parabola of the form

$$y = ax^2 + bx + c$$

where (x, y) are coordinates (n, q). The coefficients a, b, and c can be determined by solving the following three equations:

$$q_{i-1} = an_{i-1}^2 + bn_{i-1} + c.$$

$$q_i = an_i^2 + bn_i + c.$$

$$q_{i+1} = an_{i+1}^2 + bn_{i+1} + c.$$

Once a, b, c have been determined, it is straightforward to show that the ordinate at $x = n_i' = n_i + d(d = \pm 1)$ is

$$q_i' = an_i'^2 + bn_i' + c.$$

$$q_i' = q_i + \frac{d}{n_{i+1} - n_{i-1}} * \left[(n_i - n_{i-1} + d)\frac{q_{i+1} - q_i}{n_{i+1} - n_i} + (n_{i+1} - n_i - d)\frac{q_i - q_{i-1}}{n_i - n_{i-1}} \right].$$

This is the P^2 formula.

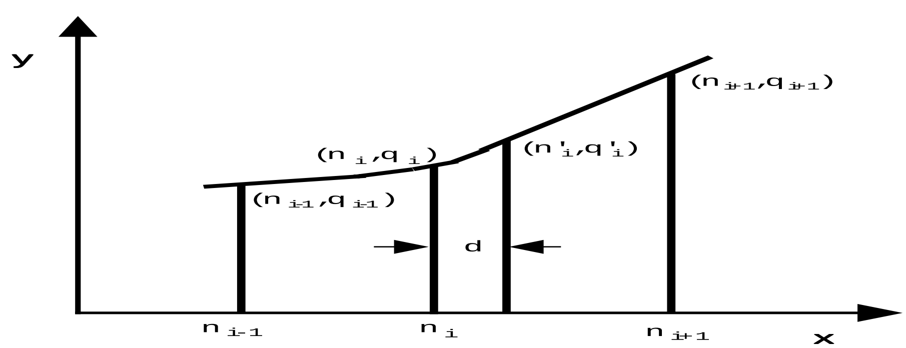

Figure 7: The P^2 formula assumes a piecewise-parabolic curve passing through three adjacent markers.

References

[EPM99] G. C. Ewing, K. Pawlikowski, and D. McNickle. Akaroa2: Exploiting network computing by distributing stochastic simulation. In *Australasian Computer Science Conference*, Submitted. Jan. 1999.

[Fis73] G. S. Fishmam. Statistical analysis for queueing simulation. *Management Science*, 20:363–369, 1973.

[GAM78] A. V. Gafarian, C. J. Ancker, and T. Morisaku. Evaluation of commonly used rules for detecting steady state in computer simulation. *Naval res. Logist. Quart.*, 78:511–529, 1978.

[HL84] P. Heidelberger and P. A. W. Lewis. Quantile estimation in dependent sequences. *Operations Research*, 32(1):185–209, Jan.-Feb. 1984.

[HW81] P. Heidelberger and P. D. Welch. A spectral method for confidence interval generation and run length control in simulations. *Communications of the ACM*, 25:233–245, 1981.

[Igl76] Donald L. Iglehart. Simulating stable stochastic systems, vi: Quantile estimation. *Journal of the Association for Computing Machinery*, 23(2):347–360, April 1976.

[JC85] Raj Jain and Imrich Chlamtac. The p2 algorithm for dynamic calculation of quantiles and histograms without storing observations. *Communications of the ACM*, 28(10):1076–1085, Oct. 1985.

[LMP98] J. R. Lee, D. McNickle, and K. Pawlikowski. A survey of confidence interval formulae for coverage analysis. Technical Report TR-COSC 04/98, Department of Computer Science, University of Canterbury, Christchurch, New Zealand, 1998.

[LMP99] J. R. Lee, D. McNickle, and K. Pawlikowski. Confidence interval estimators for coverage analysis in sequential steady-state simulation. In *Australasian Computer Science Conference*, Submitted. Jan. 1999.

[Moo80] L. W. Moore. *Quantile Estimation Methods in Regenerative processes*. PhD thesis, Department of Statistics, University of North Carolina, Chapel Hill, 1980.

[Paw90] K. Pawlikowski. Steady-state simulation of queueing processes: A survey of problems and solutions. *ACM Computing Surveys*, 22(2):122–170, June 1990.

[PYM94] K. Pawlikowski, V. Yau, and D. C. McNickle. Distributed and stochastic discrete-event simulation in parallel time streams. In *Proceedings of the 1994 Winter Simulation Conference*, pages 723–730, Dec. 1994.

[Raa87] Kimmo E. E. Raatikainen. Simultaneous estimation of several percentiles. *SIMULATION*, 49(4):159–164, 1987.

[Raa90] Kimmo E. E. Raatikainen. Sequential procedure for simultaneous estimation of several percentiles. *Transactions of The Society for Computer Simulation*, 7(1):21–44, 1990.

[Sei82a] Andrew F. Seila. Estimation of percentiles in discrete event simulation. *SIMULATION*, 39(6):193–200, Dec. 1982.

[Sei82b] Andrew F. Seila. A batching approach to quantile estimation in regenerative simulations. *Management Science*, 28(5):573–581, May 1982.

[SL79] C. H. Sauer and S. S. Lavenberg. Confidence intervals for queueing simulations of computer systems. *ACM Performance Evaluation Review*, 8(1-2):46–55, 1979.

[SST83] L. W. Schruben, H. Singh, and L. Tierney. Optimal tests for initialization bias in simulation output. *Operations Research*, 31:1167–1178, 1983.